BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 8: e140738
doi: 10.3897/biss.8.140738

OPEN ACCESS

Conference Abstract

Expressing Circumscription in the Taxon Concept Schema (TCS)

Niels Klazenga‡, Johan Liljeblad§

‡ Royal Botanic Gardens Victoria, Melbourne, Australia
§ Swedish University of Agricultural Sciences, Uppsala, Sweden

Corresponding author: Niels Klazenga (niels.klazenga@rbg.vic.gov.au)

Received: 01 Nov 2024 | Published: 01 Nov 2024

Citation: Klazenga N, Liljeblad J (2024) Expressing Circumscription in the Taxon Concept Schema (TCS). Biodiversity Information Science and Standards 8: e140738. https://doi.org/10.3897/biss.8.140738

Abstract

The Taxon Concept Schema (TCS, Biodiversity Information Standards (TDWG) 2005) is the TDWG standard for sharing taxonomic data. TCS has never enjoyed widespread use and most taxonomic data is exchanged using the Darwin Core (Wieczorek et al. 2012) Taxon class or non-standard terms. For the last three years, the TCS 2 Task Group has been working on a major new version, which will take TCS out of its XML Schema and convert it to a vocabulary standard of terms and definitions that does not dictate a data format and can be maintained under the TDWG Vocabulary Maintenance Standard (Vocabulary Maintenance Specification Task Group 2017). This new version (Taxon Concept Schema 2 Task Group 2024) is now ready to go out to public review.

With only 50 terms in total, 12 of which are borrowed from Darwin Core and Dublin Core (Dublin Core Metadata Initiative 2012), TCS is a small standard. There are, however, a lot of relationships, both internal and external, which means that all taxonomic data can be exchanged using these relatively few terms. The TCS TaxonConcept is equivalent to the Darwin Core Taxon and in some situations can be used in conjunction with the Taxon class, while in other situations it replaces it. The biggest difference between the TCS TaxonConcept and the Darwin Core Taxon is that in TCS, the Taxon Concept and Taxon Name are separated, while in Darwin Core the taxon name is embedded in the Taxon. This means that, when using TCS to exchange taxonomic data, one does not have to create—and assign identifiers to—the data artefacts that one would have to create when using Darwin Core to exchange the same data.

© Klazenga N, Liljeblad J. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Another important difference is that a TCS TaxonConcept must have a source (accordingTo). This is important because the same name can apply to different taxonomic groups. The missing element when using only names is the definition or circumscription of the taxonomic group. Circumscription is drawing a boundary around a taxonomic group and is by many taxonomists seen as the holy grail of taxonomic data. TCS 1 had the CharacterCircumscription and SpecimenCircumscription elements, but these have not yet been included in TCS 2, because we do not know how to implement them in a useful way and, if we are going to have circumscription in TCS, we want it to be operational.

While we are as yet unable to express circumscription in a meaningful way in TCS, we can still do everything we need to do with the mapping properties and the TaxonConceptMapping class that are included in TCS. Because scientific names have types and these types are specimens, some of these mappings can be derived from the synonymy, using the taxon concepts as sets and the taxon names as elements in these sets. This, along with nomenclatural business rules that decide which, of all the names that can be applied to a taxon concept, is the accepted name, is also the reason that name resolution and name matching mostly work. However, the mappings or name resolution thus obtained are only based on the information in the data set and there is no taxonomic data set that includes all the necessary mapping information in its synonymy.

The biggest challenge in dealing with taxonomic data is that we are often dealing with incomplete data. Using the right objects, i.e., taxon concepts rather than taxon names, and allowing additional expert knowledge in the form of taxon concept mappings in our data sets and name resolution, will go some way to resolve this problem.

Keywords

Darwin Core, taxonomic data, name resolution

Presenting author

Niels Klazenga

Presented at

SPNHC-TDWG 2024

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Biodiversity Information Standards (TDWG) (2005) Taxon Concept Schema (TCS). http://www.tdwg.org/standards/117
- Dublin Core Metadata Initiative (2012) DCMI Metadata Terms. https://www.dublincore.org/specifications/dublin-core/dcmi-terms/2012-06-14/
- Taxon Concept Schema 2 Task Group (2024) Taxon Concept Standard (TCS) Terms. https://github.com/tdwg/tcs2/tree/master/docs/tcs-terms. Accessed on: 2024-10-15.
- Vocabulary Maintenance Specification Task Group (2017) Vocabulary Maintenance Standard. Biodiversity Information Standards (TDWG). http://www.tdwg.org/standards/642
- Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, Robertson T, Vieglais D (2012) Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. PLoS ONE 7 (1). https://doi.org/10.1371/journal.pone.0029715